

Moving atomic-resolution imaging into the age of deep data

Oleg S. Ovchinnikov¹, S. Jesse², A. Belianinov², S. T. Pantelides^{1,2}, S.V. Kalinin², A.Y. Borisevich²

¹Department of Physics and Astronomy, Vanderbilt University, Nashville TN 37235 USA

²Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA

The current ability of electron microscopy to collect frames in ultra-quick succession allows for the examination of the material structure in a dynamic manner offering unique insights into what determines a material's properties. The ability of computer simulations based on density functional theory (DFT) to predict materials behavior has greatly increased as access to super computers has increased, making running simulations with many different starting conditions possible in a short amount of time. The combination of these two advances in fundamental science has been used on a case by case basis, which does not fully exploit the potential advantages. An approach that combines these advantages in a more general manner, allowing each to build off of the other, is an important step into the future. This can be accomplished through the use of a deep data approach that allows simulations to quickly search and analyze data for starting conditions to simulate theory and for experiments to quickly search completed simulations to pull out the underlying physics.

A major bottleneck in this approach is the need to analyze and interpret multiple frames. The process of finding all of the atoms in a frame and running analysis on the frame is a lengthy process when done manually on a single frame, let alone a movie, in which the frames must also be co-registered together. Many current analysis methods are not scalable with the ability to acquire data. This causes a large amount of collected data to go unanalyzed and unused when it might contain important physical information. In order for a deep data approach to be fully utilized, all data must be analyzed and stored in an easily searchable manner.

In this presentation, I will describe a method for atom-finding analysis of scanning transmission electron microscope (STEM) images that will allow a user to quickly analyze a single frame and then use this frame in order to analyze multi-frame movies. This will be accomplished through the use of a repository of filtering, atom finding, and data analyses (Figure 1). This ability, combined with central tracking of all operations performed on an image, will allow for both batch processing and databasing. Another advantage of such a program is to have a centralized ability to analyze atom positions and shapes using methods that have been shown to pull information out of images that would be impossible to detect with the human eye. Such methods include looking at the spatial distribution of principal component analysis (PCA) eigenvalues of the atom shape and spatial distribution of nearest neighbors (Figure 2).[1,2] Once analyzed in a uniform manner, all data can then be centrally stored for searching. Such searches could include comparing to simulations, finding similar physics across multiple materials, or any other number of reasons. Information analyses like these have great benefits for simulations such as DFT. An example of this would be using the information from the PCA of the nearest neighbors of many images of a material to narrow down the range of the initial conditions for the simulations. In order to accomplish this, ground work has been laid in the form of a test program that will combine many methods of filtering, atom finding, and data analyses into a single graphical user interface that also allows tracking of users' actions, allowing for repeatability in data analysis. This allows for data analysis to be performed on large amounts of data in a fraction of the time. This in turn allows for the ability to build a catalog of data for use in deep data analysis [3].

References

- [1] A. Borisevich, O.Ovchinnikov et al., ACS Nano **4**, 6071-6079 (2010)
 [2] A. Belianinov et al., these proceedings
 [3] The research is sponsored by the ORNL's Laboratory Directed R&D Fund (OSO, SJ, SVK, AYW) and the Division of Materials Sciences and Engineering, Office of Basic Energy Sciences, U.S. Department of Energy under the grant DE-FG02-09ER46554 (STP).

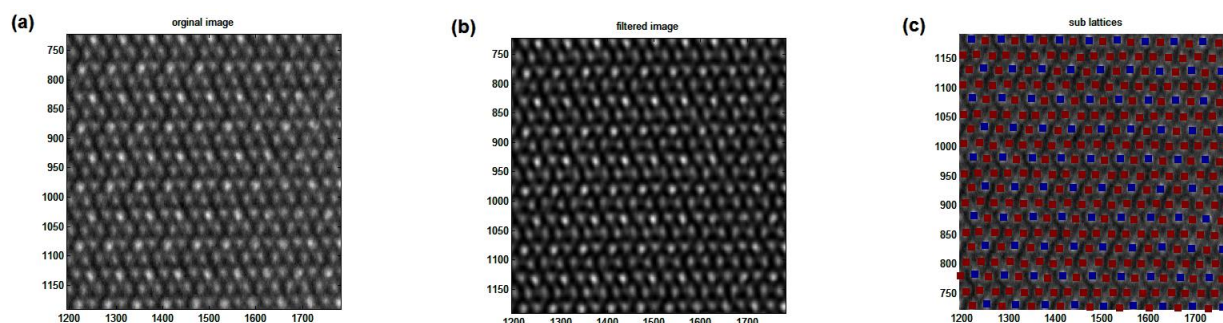


FIG 1. Example of ability to filter image, identify atom center, and filter into sub-lattices (a) original image (b) filtered with blind deconvolution, background removal, and Gaussian filter of FFT space (c) centers found and separated into two sub lattices with manually seeded points.

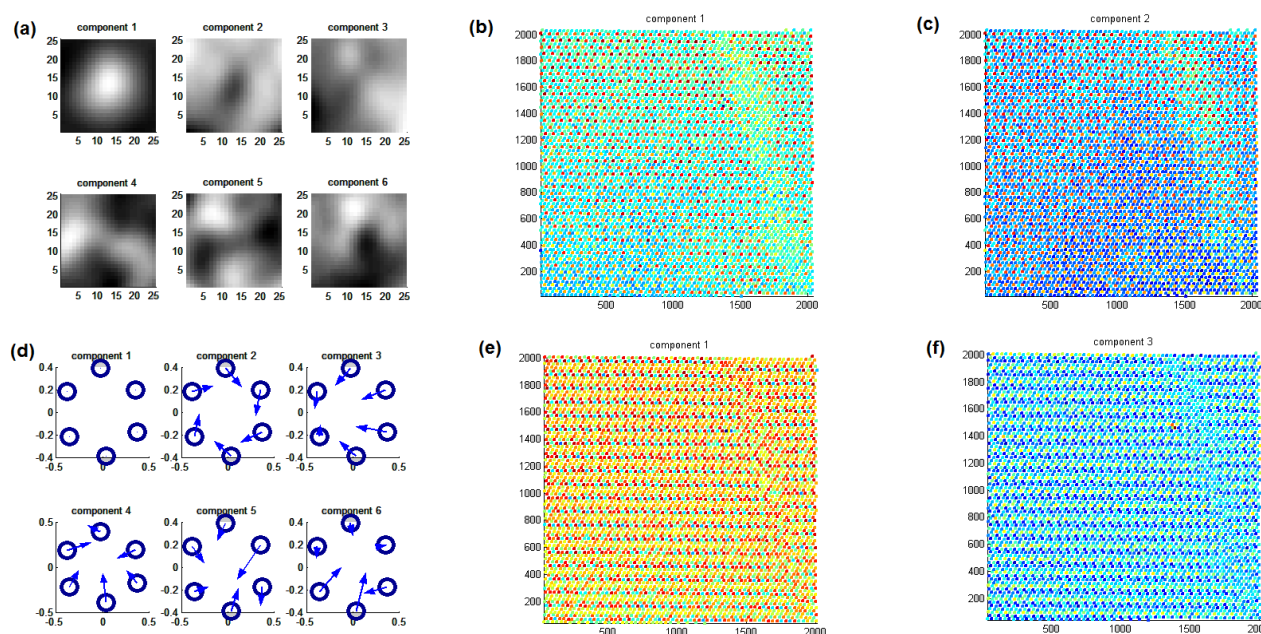


FIG 2. Example of analysis on data set (a) first 6 Eigen-vectors of atom shapes (b) spatial distribution of first Eigen-vector of atom shape showing large defect and sub lattices (c) spatial distribution of second Eigen-vector of atom shape showing large defect and sub lattices (d) first 6 Eigen-vectors of 6 nearest neighbors with 0.4 scaling factor for length (e) spatial distribution of first Eigen-vector of nearest neighbors showing large defect and sub lattices (f) spatial distribution of third Eigen-vector of nearest neighbors showing large defect and sub lattices, second Eigen-vector skipped due to showing little information.